Multi-View Fuzzy Clustering Algorithms for Multi-View Data

Kristina P. Sinaga

Chung Yuan Christian University
Department of Applied Mathematics
No. 200, Zhongbei Rd., Zhongli Dist., Taoyuan City 320314
Taiwan (R.O.C.)

June 2, 2020

中原大學
Chung Yuan Christian University

Introduction

Motivations for this Dissertation

- In many real-world problems, it is often the case that data are collected from multiple sources in diverse domains or described by various feature components.
- Processing numerous sources of data on the single-view clustering algorithm is not the right choice.
- Processing data from multiple sources is beneficial for getting more relevant information comparing to single-source data.
- Most of multi-view FCM considers view-weights instead of feature-view weights.

中原大學 Chung Yuan Christian University

Motivations for this Dissertation

To discover the structure of multi-view (multi-represented) data, we simultaneously consider the feature-view weighted and collaborative learning during the clustering (structuring) process for feature-reduction behavior in multi-view fuzzy c-means.

Objectives

Point 1

To present different types of multi-view applications

Point 2

To review the literature in the area of multi-view fuzzy clustering (MVFCM)

Point 3

To present a novel MVFCM procedure based on feature-weighting.

中原大學 Chung Yuan Christian University

Objectives Part II

Point 4

To present a novel MVFCM based on collaborative feature-weighted learning

Point 5

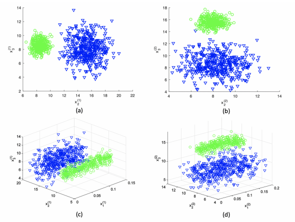
To incorporate a feature reduction behavior by considering a threshold within the clustering processes

中原大學 Chung Yuan Christian University

Multi-view Data

Example 1

An artificial spherical-shape of a two-view numerical data set with 2 clusters and 2 feature components is considered.

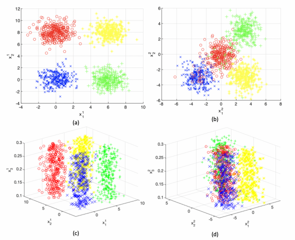


- 55% of the data points are in the blue group and 45% in the green group
- In view 2, the blue and the green ones are blended, which is it can end up to the wrong conclusion if the algorithms failed in getting the critical insight for structuring the data

中原大學
Chung Yuan Christian University

Another artificial example

An artificial spherical-shape of a two-view numerical data set with 4 clusters and 2 feature components is considered.



- Another feature generated from a uniform distribution makes the data in (a)-(b) becomes severe
- The presence of noise features in the data (a)-(b) makes the visual interpretation of a two-view numerical data set with 4 clusters chaos

中原大學
Chung Yuan Christian University

Application: UCI Machine Learning Repository

Image Segmentation

The Image segmentation (IS) data set contains 2310 instances of seven outdoor images. These seven outdoor images are Brick-face, Sky, Foliage, Cement, Window, Path, and Grass. Each instance is represented by two different views: 9 features of the shape information and 10 features of the RGB color [27 98].

中原大學 Chung Yuan Christian University

Application: MSRC data set

Figure: Sample images from the MSRC-V1 data base.

Application: MSRC data set

- The Microsoft Research Cambridge Volume 1 (MSRC-V1) data set has eight classes, and each class has 30 images [WJ05]
- We selected seven classes of eight classes with 210 images in total
- Each image is described by five views (Color moment, CENTRIST, GIST, Histogram of oriented gradients (HOG), Local binary patterns (LBP))

Application: Caltech-101 data set

Figure: Sample images from the Caltech 101-7 data base

Application: Caltech-101 data set

- Caltech 101 consists of a total of 9.146 images belonging to 101 categories [FFP07]
- We selected seven classes of 101 categories with 1474 images in total
- Each image is described by six views (Gabor, CENTRIST, Histogram of oriented gradients (HOG), Wavelet moments, Local binary patterns (LBP), GIST)

Application: Biological Data

Prokaryotic phyla

This biological data involved 551 types of Prokaryotic species where each species described by three views, including one textual data view and two genomic data views. The 2 genomic data views are known as gene repertoire and proteome composition [Brb+16].

Text Data: UCI Machine Learning Repository

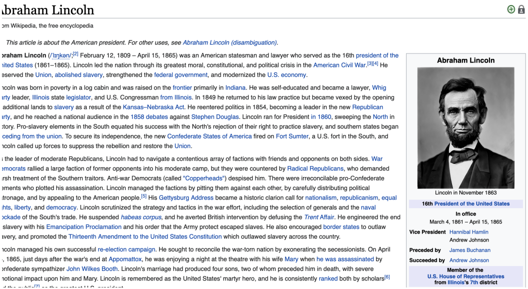
Text recognition is the process of detecting text in images and video streams and recognizing the text contained therein.

Internet Advertisements

This data set represent a set of possible advertisements on internet pages. This data set consists of 1559 columns and 3279 number of instances in the data. Each instance in the data represents one image that is tagged as an advertisement ("ad") and non-advertisement ("non-ad") in the last column [27 98]. Each instance is described by six different views (Geometry of the image, base URL, image URL, alt text, anchor text, target URL).

中原大學 Chung Yuan Christian University

Application: Multimedia data set


ıbraham Lincoln

ɔm Wikipedia, the free encyclopedia

This article is about the American president. For other uses, see Abraham Lincoln (disambiguation).

ɔraham Lincoln (/ˈlɪŋkən/;[2] February 12, 1809 – April 15, 1865) was an American statesman and lawyer who served as the 16th president of the ıited States (1861–1865). Lincoln led the nation through its greatest moral, constitutional, and political crisis in the American Civil War.[3][4] He ɘserved the Union, abolished slavery, strengthened the federal government, and modernized the U.S. economy.

ıcoln was born in poverty in a log cabin and was raised on the frontier primarily in Indiana. He was self-educated and became a lawyer, Whig ırty leader, Illinois state legislator, and U.S. Congressman from Illinois. In 1849 he returned to his law practice but became vexed by the opening additional lands to slavery as a result of the Kansas–Nebraska Act. He reentered politics in 1854, becoming a leader in the new Republican ırty, and he reached a national audience in the 1858 debates against Stephen Douglas. Lincoln ran for President in 1860, sweeping the North in ꜀tory. Pro-slavery elements in the South equated his success with the North's rejection of their right to practice slavery, and southern states began ceding from the union. To secure its independence, the new Confederate States of America fired on Fort Sumter, a U.S. fort in the South, and ıcoln called up forces to suppress the rebellion and restore the Union.

; the leader of moderate Republicans, Lincoln had to navigate a contentious array of factions with friends and opponents on both sides. War ɘmocrats rallied a large faction of former opponents into his moderate camp, but they were countered by Radical Republicans, who demanded ırsh treatment of the Southern traitors. Anti-war Democrats (called "Copperheads") despised him. There were irreconcilable pro-Confederate ɘments who plotted his assassination. Lincoln managed the factions by pitting them against each other, by carefully distributing political ıtronage, and by appealing to the American people.[5] His Gettysburg Address became a historic clarion call for nationalism, republicanism, equal ıhts, liberty, and democracy. Lincoln scrutinized the strategy and tactics in the war effort, including the selection of generals and the naval ɔckade of the South's trade. He suspended *habeas corpus*, and he averted British intervention by defusing the *Trent* Affair. He engineered the end slavery with his Emancipation Proclamation and his order that the Army protect escaped slaves. He also encouraged border states to outlaw ıvery, and promoted the Thirteenth Amendment to the United States Constitution which outlawed slavery across the country.

ıcoln managed his own successful re-election campaign. He sought to reconcile the war-torn nation by exonerating the secessionists. On April ·, 1865, just days after the war's end at Appomattox, he was enjoying a night at the theatre with his wife Mary when he was assassinated by ɔnfederate sympathizer John Wilkes Booth. Lincoln's marriage had produced four sons, two of whom preceded him in death, with severe ıotional impact upon him and Mary. Lincoln is remembered as the United States' martyr hero, and he is consistently ranked both by scholars[6]

Abraham Lincoln

Lincoln in November 1863

16th President of the United States

In office

March 4, 1861 – April 15, 1865

Vice President Hannibal Hamlin
Andrew Johnson

Preceded by James Buchanan

Succeeded by Andrew Johnson

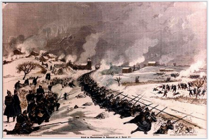
Member of the U.S. House of Representatives from Illinois's 7th district

Figure: Example of a combination of two data types such as text and images: section from the Wikipedia article on the Abraham Lincoln

Application: Multimedia data set II

Industrial warfare [edit source]

Further information: Industrial warfare



Franco-Prussian War

As weapons—particularly small arms—became easier to use, countries began to abandon a complete reliance on professional soldiers in favor of conscription. Technological advances became increasingly important; while the armies of the previous period had usually had similar weapons, the industrial age saw encounters such as the Battle of Sadowa, in which possession of a more advanced technology played a decisive role in the outcome.[69] Conscription was employed in industrial warfare to increase the number of military personnel that were available for combat. Conscription was notably used by Napoleon Bonaparte and the major parties during the two World Wars.

Figure: Example of a combination of two data types such as text and images: part of a series on War history (Industrial warfare category).

Wikipedia Articles: Figures results

- The value of retrieval of some text and images gathers a wealth of historical information about Abraham Lincoln and Industrial warfare
- The representation of information and knowledge becomes more explorable and understandable to the human users who need insights on specific topic

Multimedia: Wikipedia Articles

- Costa Pereira et al. retrieved 2669 articles, and each article is assigned over 29 categories, including the wiki text and image [Cos+14].
- Ten top-level categories are highlighted such as art architecture, geography places, history, literature theatre, biology, media, music, sport recreation, royalty nobility, and warfare.
- A random split was used to produce a training set of 2173 documents and a test set of 693 documents.

Data preparation process

WHEN
WHY
HOW

Data preparation process: WHEN?

The high dimensional data sets are often sparse, heterogeneous, uncertain, incomplete, missing, and derive.

中原大學 Chung Yuan Christian University

Data preparation process: WHY?

- There is a need for a unified and coherent machine learning framework that can analyze Big Data of different data characteristics and different application problem domains
- Each machine learning models has different circumstances in discovering the structured of data
- The accuracy of the algorithms is largely influenced by the type of data characteristics.
- To fix the failures of data clustering algorithms under different types of variations in the data

中原大學 Chung Yuan Christian University

Data preparation process: HOW?

1. Data cleaning
2. Data integration
3. Data transformation
4. Data reduction
5. Data discretization

Clustering Algorithms

K-means clustering

Objective function

$$J(U,A)=\sum_{i=1}^{n}\sum_{k=1}^{c}\mu_{ik}\|x_i-a_k\|^2 \tag{1}$$

Update equations

$$a_k=\frac{\sum_{i=1}^{n}\mu_{ik}x_{ij}}{\sum_{i=1}^{n}\mu_{ik}} \tag{2}$$

$$\mu_{ik}=\begin{cases}1 & \text{if } \|x_i-a_k\|^2=\min\limits_{1\le k\le c}\|x_i-a_k\|^2\\ 0 & \text{Otherwise}\end{cases} \tag{3}$$

中原大學 Chung Yuan Christian University

Fuzzy c-means clustering

Objective function

$$J_{FCM}(U, A) = \sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^{m} \|x_i - a_k\|^2 \tag{4}$$

Update equations

$$a_k = \frac{\sum_{i=1}^{n} \mu_{ik}^{m} x_{ij}}{\sum_{i=1}^{n} \mu_{ik}^{m}} \tag{5}$$

$$\mu_{ik} = \left(\sum_{k'=1}^{c} (d_{ik}/d_{ik'})^{\frac{2}{m-1}} \right)^{-1} \tag{6}$$

中原大學 Chung Yuan Christian University

Co-FC (Collaborative Fuzzy c-means clustering)

Objective function

$$J_{Co-FC}\left(h\right)=\sum_{i=1}^{n}\sum_{k=1}^{c}\left(\mu_{ik}^{h}\right)^{m}\sum_{j=1}^{d_h}\left(x_{ij}^{h}-a_{kj}^{h}\right)^{2} + \sum_{h'=1,h'\neq h}^{s}\alpha_{hh'}\sum_{i=1}^{n}\sum_{k=1}^{c}\left(\mu_{ik}^{h}-\mu_{ik}^{h'}\right)^{2}\sum_{j=1}^{d_h}\left(x_{ij}^{h}-a_{kj}^{h}\right)^{2} \tag{7}$$

where $\alpha_{hh'}$ is a disagreement values between the hth database and the $h'th$ database [Ped02].

中原大學 Chung Yuan Christian University

Co-FC: Limitations

- This technique does not applicable for multi-sources data
- It does not clear how to estimate an appropriate values of $\alpha_{hh'}$ such as different data plays different roles of $\alpha_{hh'}$.

中原大學 Chung Yuan Christian University

Co-FKM: Objective function

Objective function [Cle+09]

$$J_{Co-FKM} = \sum_{h=1}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \left(\mu_{ik}^{h}\right)^{m} \sum_{j=1}^{d_h} \left(x_{ij}^{h} - a_{kj}^{h}\right)^{2} + \eta\Delta \qquad (8)$$

where

$$\Delta = \frac{1}{s-1} \sum_{h'=1, h' \neq h}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \left\{ \left(\mu_{ik}^{h'}\right)^{m} - \left(\mu_{ik}^{h}\right)^{m} \right\} \sum_{j=1}^{d_h} \left(x_{ij}^{h} - a_{kj}^{h}\right)^{2}$$

Co-FKM: Limitations

- Treated the importance of each view equally
- Parameter η is sensitive and user define
- The fuzzier parameter m is sensitive

WV-Co-FCM: Objective function

Objective function [Jia+15]

$$J_{WV-Co-FCM} = \sum_{h=1}^{s} v_h \left(J_{FCM}^{h} + \Delta_h \right) + \beta \sum_{h=1}^{s} v_h \ln v_h \tag{9}$$

where

$$J_{FCM}^{h} = \sum_{i=1}^{n} \sum_{k=1}^{c} \left(\mu_{ik}^{h} \right)^{m} \sum_{j=1}^{d_h} \left(x_{ij}^{h} - a_{kj}^{h} \right)^{2}$$

中原大學 Chung Yuan Christian University

WV-Co-FCM: Short Explanation

$$\Delta_h = \sum_{i=1}^{n} \alpha_{ik}^h \sum_{k=1}^{c} \mu_{ik}^h \left(1 - \left(\mu_{ik}^h\right)^{m-1}\right) - \sum_{i=1}^{n} \delta_{ik}^h \sum_{k=1}^{c} \mu_{ik}^h \left(1 - \left(\mu_{ik}^h\right)^{m-1}\right)$$

$\delta_{\mathbf{ik}}^{\mathbf{h}}$ has four cases

Case 1: $\eta \frac{1}{s-1} \sum_{h'=1, h'\neq h}^{s} \sum_{j=1}^{d_{h'}} \left(x_{ij}^{h'} - a_{kj}^{h'}\right)^2$

Case 2: $\frac{\eta}{s} \sum_{h=1}^{s} \sum_{j=1}^{d_h} \left(x_{ij}^{h} - a_{kj}^{h}\right)^2$

Case 3: $\eta \min\limits_{h'\neq h} \left\{ \sum_{j=1}^{d_{h'}} \left(x_{ij}^{h'} - a_{kj}^{h'}\right)^2 \right\}$

Case 4: $\eta \sqrt[s-1]{\prod_{h'\neq h} \sum_{j=1}^{d_{h'}} \left(x_{ij}^{h'} - a_{kj}^{h'}\right)^2}$

While

$$\alpha_{ik}^h = \eta \sum_{j=1}^{d_h} \left(x_{ij}^h - a_{kj}^h\right)^2$$

中原大學 Chung Yuan Christian University

WV-Co-FCM: Limitations

- Treated the importance of feature-view equally
- Parameter β and η are sensitive
- The fuzzier parameter m is sensitive

Global Solutions: Co-FKM & WV-Co-FCM

- **Co-FKM**

$$\tilde{\mu}_{ik} = \sqrt[s]{\prod_{h=1}^{s} \mu_{ik}^{h}} \tag{10}$$

- **WV-Co-FCM**

$$\tilde{U} = \sum_{h=1}^{s} v_h U^h \tag{11}$$

Minimax-FCM

Objective function

$$J_{\text{minimax-FCM}}\left(V, U^*, A^h\right) = \sum_{h=1}^{s} (v_h)^{\beta} \sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^*)^m \sum_{j=1}^{d_h} \left(x_{ij}^h - a_{kj}^h\right)^2 \tag{12}$$

In the minimax-FCM, Wang and Chen [WC17] used the min-max optimization with

$$\min_{U^*, \{A^h\}_{h=1}^s} \max_{\{v_h\}_{h=1}^s} \sum_{h=1}^{s} (v_h)^{\beta} \sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^*)^m \sum_{j=1}^{d_h} \left(x_{ij}^h - a_{kj}^h\right)^2$$

Minimax-FCM: Limitations

- Treated the importance of feature-view equally
- Parameter β is a user define
- The fuzzier parameter m is sensitive

中原大學 Chung Yuan Christian University

MultiNMF: Objective Function

Objective function

$$J_{\text{MultiNMF}} = \sum_{h=1}^{s} \left\| X^h - U^h \left(A^h \right)^T \right\|_F^2 + \sum_{h=1}^{s} \lambda_h D\left(A^h, A^* \right) \quad (13)$$

- Where $D\left(A^h, A^*\right) = \left\| A^h - A^* \right\|_F^2$ is the Frobenius norm and used as a measure of disagreement between the cluster indicator matrix A^h and the cluster consensus matrix A^* of the hth view [Liu+13].

中原大學 Chung Yuan Christian University

MultiNMF: Limitations

- Treated the importance of each view equally
- Treated the importance of feature-view equally
- Parameter λ_h is a user define and sensitive
- Depends on normalization process because Input data must be in a positive value

中原大學 Chung Yuan Christian University

The Proposed Algorithms

Feature-weighted MVFCM (FW-MVFCM)

FW-MVFCM: Introduction

- The FW-MVFCM deal with multi-representated data
- The goal: Compute a similarity matrix for data object/point which appear in different views by considering the importance of each feature-view components.
- The idea: create a simple feature weighting in multi-view FCM by incorporated the local membership of different views directly into a single global membership.

FW-MVFCM: Objective function

Objective function

$$J_{FW-MVFCM}=\sum_{h=1}^{s}(v_h)^{\beta}\sum_{i=1}^{n}\sum_{k=1}^{c}\mu_{ik}^{m}\sum_{j=1}^{d_h}(w_j^h)^2\left(x_{ij}^h-a_{kj}^h\right)^2 \tag{14}$$

Subject to

$$\begin{aligned}&\sum_{k=1}^{c}\mu_{ik}=1,\mu_{ik}\in[0,1]\\&\sum_{h=1}^{s}v_h=1,v_h\in[0,1]\\&\sum_{j=1}^{d_h}w_j^h=1,w_j^h\in[0,1].\end{aligned} \tag{15}$$

中原大學 Chung Yuan Christian University

FW-MVFCM: Theorem 1

The updating equations for necessary conditions to minimize the $J_{FW-MVFCM}$ of Equations (14) are

$$a_{kj}^{h} = \frac{\sum_{i=1}^{n} \mu_{ik}^{m} x_{ij}^{h}}{\sum_{i=1}^{n} \mu_{ik}^{m}} \tag{16}$$

$$v_h = \left(\sum_{r=1}^{s} \left(\frac{\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^{m} (w_j^h)^2 \left(x_{ij}^{h} - a_{kj}^{h} \right)^2}{\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^{m} (w_j^r)^2 \left(x_{ij}^{r} - a_{kj}^{r} \right)^2} \right)^{\frac{1}{\beta-1}} \right)^{-1} \tag{17}$$

中原大學 Chung Yuan Christian University

FW-MVFCM: Theorem 1 Part II

$$w_j^h = \sum_{j'=1}^{d_h} \left(\frac{(v_h)^\beta \sum_{i=1}^n \sum_{k=1}^c \mu_{ik}^m \left(x_{ij}^h - a_{kj}^h \right)^2}{(v_h)^\beta \sum_{i=1}^n \sum_{k=1}^c \mu_{ik}^m \left(x_{ij'}^h - a_{kj'}^h \right)^2} \right)^{-1} \tag{18}$$

$$\mu_{ik} = \left(\left(\sum_{l=1}^{c} \frac{\sum_{h=1}^s (v_h)^\beta \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{kj}^h \right)^2}{\sum_{h=1}^s (v_h)^\beta \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{lj}^h \right)^2} \right)^{\left(\frac{1}{m-1}\right)} \right)^{-1} \tag{19}$$

FW-MVFCM: Proof μ_{ik}

We prove it by using Langrange multipliers. The Lagrangian of $J_{FW-MVFCM}$ subject to $\sum_{k=1}^{c} \mu_{ik} = 1, \mu_{ik} \in [0,1]$ is
$\hat{J} = J_{FW-MVFCM} - \lambda_1 \left(\sum_{k=1}^{c} \mu_{ik} - 1\right)$

$$\frac{\partial \tilde{J}}{\partial \mu_{ik}} = m\mu_{ik}^{m-1} \sum_{h=1}^{s} (v_h)^{\beta} \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{kj}^h\right)^2 - \lambda_1 = 0 \quad (20)$$

$$\mu_{ik} = \left(\frac{\lambda_1}{\sum_{h=1}^{s} (v_h)^{\beta} \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{kj}^h\right)^2} \right)^{\frac{1}{m-1}} \quad (21)$$

FW-MVFCM: Proof μ_{ik}

- Since $\sum_{l=1}^{c} \mu_{il} = 1$ we get

$$\lambda_1 = \frac{1}{\sum_{l=1}^{c} \left(\frac{1}{\sum_{h=1}^{s} (v_h)^{\beta} \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{lj}^h \right)^2} \right)^{\frac{1}{m-1}}} \tag{22}$$

- Thus, the updating equation for μ_{ik} can be obtained as Equation (19)

中原大學 Chung Yuan Christian University

FW-MVFCM: Proof v_h

- We consider the Lagrangian of (14) subject to $\sum_{h=1}^{s} v_h = 1$ with $\hat{J} = J_{FW-MVFCM} - \lambda_2 \left(\sum_{h=1}^{s} v_h - 1\right)$

$$\frac{\partial \tilde{J}}{\partial v_h} = \beta (v_h)^{\beta-1} \left(\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^{m} \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{kj}^h \right)^2 \right) - \lambda_2 = 0 \quad (23)$$

$$v_h = (\lambda_2)^{\frac{1}{\beta-1}} \left(\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^{m} \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^h - a_{kj}^h \right)^2 \right)^{\frac{-1}{\beta-1}} \quad (24)$$

FW-MVFCM: Proof v_h

- Since $\sum_{r=1}^{s} v_r = 1$ we get

$$(\lambda_2)^{\frac{1}{\beta-1}} = \left(\sum_{r=1}^{s} \left(\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^{m} \sum_{j=1}^{d_r} (w_j^r)^2 \left(x_{ij}^r - a_{kj}^r \right)^2 \right)^{\frac{-1}{\beta-1}} \right)^{-1} \tag{25}$$

- Thus, the updating equation for v_h can be obtained as Equation (17)

中原大學 Chung Yuan Christian University

FW-MVFCM: Proof w_j^h

- We finally consider the Lagrangian of (14) subject to $\sum_{j=1}^{d_h} w_j^h = 1$ with $\hat{J} = J_{FW-MVFCM} - \lambda_3 \left(\sum_{j=1}^{d_h} w_j^h - 1 \right)$

$$\frac{\partial \tilde{J}}{\partial w_j^h} = (v_h)^\beta \left(2 \sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^m w_j^h \left(x_{ij}^h - a_{kj}^h \right)^2 \right) - \lambda_3 = 0 \tag{26}$$

$$w_j^h = \lambda_3 \left((v_h)^\beta \left(\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^m \left(x_{ij}^h - a_{kj}^h \right)^2 \right) \right)^{-1} \tag{27}$$

中原大學 Chung Yuan Christian University

FW-MVFCM: Proof w_j^h

- Since $\sum_{j'=1}^{d_h} w_{j'}^h = 1$ we get

$$\lambda_3 = \left(\sum_{j'=1}^{d_h} \left((v_h)^\beta \left(\sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik}^m \left(x_{ij'}^h - a_{kj'}^h \right)^2 \right) \right)^{-1} \right)^{-1} \tag{28}$$

- Thus, the updating equation for w_j^h can be obtained as Equation (18)

中原大學 Chung Yuan Christian University

How FW-MVFCM's Algorithm Works

So how does a FW-MVFCM Algorithm work, I will explain this using six simple steps:

Input Data set $\boldsymbol{X} = \{x_1, ..., x_n\}$ with $x_i^h = \{x_{ij}^h\}_{j=1}^{d_h}$, number c of cluster, $\beta = 0$, and $\varepsilon > 0$

Output $w_j^h, v_h, a_{kj}^h, \mu_{ik}$.

Initialization Randomly generate initial U, initialize feature weight $W^{h^t} = [w_j^h]_{1\times d_h}$ (user may define $w_j^h = 1/d_h \forall j$), initialize view weight $V^t = [v_h]_{1\times s}$ (user may define $v_h = 1/s \forall h$), and set $t = 1$.

Step 1 Calculate β by $\beta = t/n$.

How FW-MVFCM's Algorithm Works

Step 2 Udate the feature weight W^{h^t} using $A^{h^{t-1}}, U^{t-1}$, and V^{t-1} by Eq. (18) .

Step 3 Update the view weight V^{h^t} using $A^{h^{t-1}}, U^{t-1}$, and W^{h^t} by Eq. (17).

Step 4 Update the cluster centers A^{h^t} using U^{t-1} by Eq. (16).

Step 5 Update membership U^t using $W^{h^t}$and A^{h^t} by Eq. (19).

Step 6 If $\left(\left|\|U^t\| - \|U^{t-1}\|\right|\right) \Big/ nc < \varepsilon$, then stop;
Else set $t = t + 1$ and go back to Step 1.

中原大學 Chung Yuan Christian University

From FW-MVFCM to Co-FW-MVFCM

Modification of FW-MVFCM: Introduction

- The modification of FW-MVFCM's objective function is made to create the collaborative learning within multi-represantated data
- The idea: added the memberships of different views into the FW-MVFCM

Modification of FW-MVFCM: Objective Function

Based on objective function of Eq. (14)

$$J_{FW-MVFCM} = \sum_{h=1}^{s} (v_h)^{\beta} \sum_{i=1}^{n} \sum_{k=1}^{c} \left(\mu_{ik}^{h}\right)^{m} \sum_{j=1}^{d_h} (w_j^h)^2 \left(x_{ij}^{h} - a_{kj}^{h}\right)^2 \tag{29}$$

Subject to

$$\sum_{k=1}^{c} \mu_{ik}^{h} = 1, \mu_{ik}^{h} \in [0,1]$$
$$\sum_{h=1}^{s} v_h = 1, v_h \in [0,1] \tag{30}$$
$$\sum_{j=1}^{d_h} w_j^{h} = 1, w_j^{h} \in [0,1].$$

Collaborative FW-MVFCM (Co-FW-MVFCM)

Co-FW-MVFCM: Introdution

- Although we previously proposed the simplest one of feature-weighted MVFCM algorithms for clustering multi-view data. However, the correct classification rate is not enough that can be improved.
- This is because the recognition rate of the proposed feature-weighted MVFCM algorithm can evaluate the error of misclassification contained in the final partitioning, but it provides only partial information about the structural quality of the obtaining partition.
- That is, it may improve the clustering performance but the resulting partition might be practically useless. In order to get a more objective judgement about the partitioning quality, the exchanging information about the local partition between each view during the learning process should be taken into considerations.

Co-FW-MVFCM: Objective Function

By referring to Equation (7) of Pedrycz [17], we provide the collaborative learning into the objective function of Equation (29) with

$$
\begin{aligned}
J_{Co-FW-MVFCM} = & \sum_{h=1}^{s} (v_h)^{\beta} \sum_{i=1}^{n} \sum_{k=1}^{c} \left(\mu_{ik}^{h}\right)^2 \left(d_{ik,w}^{h}\right)^2 \\
& + \alpha \sum_{\substack{h'=1 \\ h' \neq h}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{h} - \mu_{ik}^{h'}\}^2 \left(d_{ik,w}^{h}\right)^2
\end{aligned}
\tag{31}
$$

$$
\text{with } d_{ik,w}^{h} = \sqrt{\sum_{j=1}^{d_h} \left(w_j^h\right)^2 \left(x_{ij}^h - a_{kj}^h\right)^2} \tag{32}
$$

Co-FW-MVFCM: Constraints

Subject to

$$\sum_{k=1}^{c} \mu_{ik}^{h} = 1, \mu_{ik}^{h} \in [0,1]$$
$$\sum_{h=1}^{s} v_{h} = 1, v_{h} \in [0,1] \tag{33}$$
$$\sum_{j=1}^{d_h} w_{j}^{h} = 1, w_{j}^{h} \in [0,1].$$

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Highlighted Terminology

Highlighted 1

The first terms of Eq. (31) is a single view partition process to produce a local clustering partition in each view.

Highlighted 2

The second terms of Eq. (31) corresponds to the collaborative learning between views in multi-view data.

Highlighted 3

$m = 2$ is a new regulation to optimize the parameters of the second terms of Eq. (31).

中原大學 Chung Yuan Christian University

Global Solutions: Co-FW-MVFCM

In order to merge the partition matrices in each view and obtain the global clustering result, we adopt the summation of each weighted fuzzy partition matrix for each view as:

$$\tilde{U} = \sum_{h=1}^{s} v_h U^h \tag{34}$$

Co-FW-MVFCM: Theorem 2

The updating equations for necessary conditions to minimize the $J_{Co-FW-MVFCM}$ of Equation (31) are

$$\mu_{ik}^{h} = \left(1 - \sum_{l=1}^{c}\left(\frac{\varphi_{il}^{h}}{1+\alpha}\right)\right)\left(\sum_{l=1}^{c}\frac{(d_{ik,w}^{h})^{2}}{(d_{il,w}^{h})^{2}}\right)^{-1} + \frac{\varphi_{ik}^{h}}{1+\alpha} \qquad (35)$$

Co-FW-MVFCM: Theorem 2 Part II

$$v_h = \left(\sum_{r=1}^{s} \left(\frac{\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{h})^2 (d_{ik,w}^{h})^2}{\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{r})^2 (d_{ik,w}^{r})^2} + \frac{\alpha \sum_{\substack{h'=1 \\ h' \neq h}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{h} - \mu_{ik}^{h'}\}^2 (d_{ik,w}^{h})^2}{\alpha \sum_{\substack{h'=1 \\ h' \neq r}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{r} - \mu_{ik}^{h'}\}^2 (d_{ik,w}^{r})^2)} \right)^{\frac{1}{\beta-1}} \right)^{-1} \tag{36}$$

Co-FW-MVFCM: Theorem 2 Part III

$$w_j^h = \left(\sum_{j'=1}^{d_h} \left(\frac{(v_h)^\beta \left(\sum_{i=1}^n \sum_{k=1}^c (\mu_{ik}^h)^2 (x_{ij}^h - a_{kj}^h)^2 \right.}{(v_h)^\beta \left(\sum_{i=1}^n \sum_{k=1}^c (\mu_{ik}^h)^2 (x_{ij'}^h - a_{kj'}^h)^2 \right.} + \frac{\left. \alpha \sum_{\substack{h'=1 \\ h' \neq h}}^s \sum_{i=1}^n \sum_{k=1}^c \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 (x_{ij}^h - a_{kj}^h)^2 \right)}{\left. \alpha \sum_{\substack{h'=1 \\ h' \neq h}}^s \sum_{i=1}^n \sum_{k=1}^c \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 (x_{ij'}^h - a_{kj'}^h)^2 \right)} \right) \right)^{-1} \tag{37}$$

Co-FW-MVFCM: Theorem 2 Part IV

$$a_{kj}^{h} = \frac{(v_h)^{\beta}\sum_{i=1}^{n}(\mu_{ik}^{h})^{2}(w_j^{h})^{2}x_{ij}^{h}}{(v_h)^{\beta}\sum_{i=1}^{n}(\mu_{ik}^{h})^{2}(w_j^{h})^{2}} + \frac{\alpha\sum_{\substack{h'=1\\h'\neq h}}^{s}(v_h)^{\beta}\sum_{i=1}^{n}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^{2}(w_j^{h})^{2}x_{ij}^{h}}{\alpha\sum_{\substack{h'=1\\h'\neq h}}^{s}(v_h)^{\beta}\sum_{i=1}^{n}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^{2}(w_j^{h})^{2}} \tag{38}$$

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

Shortened Proof:

Membership μ_{ik}^h

$$\mu_{ik}^h = \left(1 - \sum_{l=1}^{c}\left(\frac{\varphi_{il}^h}{1+\alpha}\right)\right)\left(\sum_{l=1}^{c}\frac{(d_{ik,w}^h)^2}{(d_{il,w}^h)^2}\right)^{-1} + \frac{\varphi_{ik}^h}{1+\alpha}$$

Proof (by Langrangian multipliers): The Lagrangian of $J_{Co-FW-MVFCM}$ subject to $\sum_{k=1}^{c}\mu_{ik}^h = 1, \mu_{ik}^h \in [0,1]$ is $\hat{J} = J_{Co-FW-MVFCM} - \lambda_1\left(\sum_{k=1}^{c}\mu_{ik}^h - 1\right)$

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

By taking the partial derivative of the Lagrangian w.r.t. μ_{ik}^h and setting them to be zero, we obtain the equation

$$\frac{\partial \tilde{J}}{\partial \mu_{ik}^h} = (v_h)^\beta \left(2\mu_{ik}^h (d_{ik,w}^h)^2 + 2\alpha \sum_{\substack{h'=1 \\ h' \neq h}}^{s} \{\mu_{ik}^h - \mu_{ik}^{h'}\} (d_{ik,w}^h)^2 \right) - \lambda_1 = 0 \tag{39}$$

Thus,

$$\mu_{ik}^h = \frac{\lambda_1 + (v_h)^\beta \alpha \sum_{h'=1, h' \neq h}^{s} \mu_{ik}^{h'} (d_{ik,w}^h)^2}{(v_h)^\beta (1+\alpha)(d_{ik,w}^h)^2} \tag{40}$$

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

Now, let $\varphi_{ik}^{h} = \alpha \sum_{h'=1,h'\neq h}^{s} \mu_{ik}^{h'}$. We can rewrite equation (40) becomes:

$$\mu_{ik}^{h} = \frac{\lambda_1}{(v_h)^{\beta}(1+\alpha)} \frac{1}{(d_{ik,w}^{h})^2} + \frac{\varphi_{ik}^{h}}{1+\alpha} \tag{41}$$

Since, $\sum_{l=1}^{c} \mu_{il} = 1$, we get:

$$\frac{\lambda_1}{(v_h)^{\beta}(1+\alpha)} = \frac{1 - \sum_{l=1}^{c} \left(\frac{\varphi_{il}^{h}}{1+\alpha}\right)}{\sum_{l=1}^{c} \frac{1}{(d_{il,w}^{h})^2}} \tag{42}$$

Thus, the updating equation for μ_{ik}^{h} can be obtained as Equation (35)

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

Shortened Proof:

View weight v_h

$$v_h = \left(\sum_{r=1}^{s} \left(\frac{\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{h})^2 (d_{ik,w}^{h})^2}{\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{r})^2 (d_{ik,w}^{r})^2} + \frac{\alpha \sum_{\substack{h'=1 \\ h' \neq h}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{h} - \mu_{ik}^{h'}\}^2 (d_{ik,w}^{h})^2}{\alpha \sum_{\substack{h'=1 \\ h' \neq r}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{r} - \mu_{ik}^{h'}\}^2 (d_{ik,w}^{r})^2)} \right)^{\frac{1}{\beta-1}} \right)^{-1}$$

Proof The Lagrangian of $J_{Co-FW-MVFCM}$ subject to $\sum_{h=1}^{s} v_h = 1$ is $\hat{J} = J_{Co-FW-MVFCM} - \lambda_2 \left(\sum_{h=1}^{s} v_h - 1\right)$

Co-FW-MVFCM: Shortened Proof

By taking the partial derivative of the Lagrangian w.r.t. v_h and setting them to be zero, we obtain the equation

$$
\begin{aligned}
\frac{\partial \tilde{J}}{\partial v_h} &= \beta (v_h)^{\beta-1} \Big(\textstyle\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{h})^2 (d_{ik,w}^{h})^2 \\
&\quad + \alpha \textstyle\sum_{h'=1,h'\neq h}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{h} - \mu_{ik}^{h'}\}^2 (d_{ik,w}^{h})^2 \Big) - \lambda_2 = 0
\end{aligned}
$$

$$
\begin{aligned}
v_h &= (\lambda_2)^{\frac{1}{\beta-1}} \Big(\textstyle\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{h})^2 (d_{ik,w}^{h})^2 \\
&\quad + \alpha \textstyle\sum_{h'=1,h'\neq h}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{h} - \mu_{ik}^{h'}\}^2 (d_{ik,w}^{h})^2 \Big)^{\frac{-1}{\beta-1}}
\end{aligned}
$$

中原大學
Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

Since, $\sum_{r=1}^{s} v_r = 1$, we get:

$$(\lambda_2)^{\frac{1}{\beta-1}} = \Bigg(\sum_{r=1}^{s} \Big(\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^{r})^2 (d_{ik,w}^{r})^2 + \alpha \sum_{h'=1,h'\neq r}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^{r} - \mu_{ik}^{h'}\} (d_{ik,w}^{r})^2 \Big)^{\frac{-1}{\beta-1}} \Bigg)^{-1}$$

Thus, the updating equation for v_h can be obtained as Equation (36)

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

Shortened Proof:

Feature weight w_j^h

$$w_j^h = \left(\sum_{j'=1}^{d_h} \left(\frac{(v_h)^{\beta} \left(\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^h)^2 (x_{ij}^h - a_{kj}^h)^2 \right.}{(v_h)^{\beta} \left(\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^h)^2 (x_{ij'}^h - a_{kj'}^h)^2 \right.} + \frac{\alpha \sum_{\substack{h'=1 \\ h' \neq h}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 (x_{ij}^h - a_{kj}^h)^2 \Big)}{\alpha \sum_{\substack{h'=1 \\ h' \neq h}}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 (x_{ij'}^h - a_{kj'}^h)^2 \Big)} \right) \right)^{-1}$$

Proof The Lagrangian of $J_{Co-FW-MVFCM}$ subject to $\sum_{j=1}^{d_h} w_j^h = 1$ is $\hat{J} = J_{Co-FW-MVFCM} - \lambda_3 \left(\sum_{j=1}^{d_h} w_j^h - 1 \right)$

Co-FW-MVFCM: Shortened Proof

By taking the partial derivative of the Lagrangian w.r.t. w_j^h and setting them to be zero, we obtain the equation

$$\begin{aligned}\frac{\partial \tilde{J}}{\partial w_j^h} &= (v_h)^\beta \Big(2\sum_{i=1}^n \sum_{k=1}^c (\mu_{ik}^h)^2 w_j^h (x_{ij}^h - a_{kj}^h)^2 \\ &\quad + 2\alpha \sum_{h'=1,h'\neq h}^s \sum_{i=1}^n \sum_{k=1}^c \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 w_j^h (x_{ij}^h - a_{kj}^h)^2\Big) - \lambda_3 = 0\end{aligned}$$

Thus, we have

$$\begin{aligned}w_j^h &= \lambda_3 \Bigg((v_h)^\beta \Big(\sum_{i=1}^n \sum_{k=1}^c (\mu_{ik}^h)^2 (x_{ij}^h - a_{kj}^h)^2 \\ &\quad + \alpha \sum_{h'=1,h'\neq h}^s \sum_{i=1}^n \sum_{k=1}^c \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 (x_{ij}^h - a_{kj}^h)^2\Big)\Bigg)^{-1}\end{aligned}$$

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Shortened Proof

Since, $\sum_{j'=1}^{d_h} w_{j'}^h = 1$, we get:

$$\lambda_3 = \Bigg(\sum_{j'=1}^{d_h} \Big((v_h)^\beta \big(\sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik}^h)^2 (x_{ij'}^h - a_{kj'}^h)^2 + \alpha \sum_{h'=1, h' \neq h}^{s} \sum_{i=1}^{n} \sum_{k=1}^{c} \{\mu_{ik}^h - \mu_{ik}^{h'}\}^2 (x_{ij'}^h - a_{kj'}^h)^2 \big) \Big)^{-1} \Bigg)^{-1}$$

Thus, the updating equation for w_j^h can be obtained as Equation (37)

Co-FW-MVFCM: Shortened Proof

Shortened Proof:

Cluster centers a_{kj}^{h}

$$a_{kj}^{h}=\frac{(v_h)^{\beta}\sum_{i=1}^{n}(\mu_{ik}^{h})^{2}(w_j^{h})^{2}x_{ij}^{h}+\alpha\sum_{\substack{h'=1\\h'\neq h}}^{s}(v_h)^{\beta}\sum_{i=1}^{n}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^{2}(w_j^{h})^{2}x_{ij}^{h}}{(v_h)^{\beta}\sum_{i=1}^{n}(\mu_{ik}^{h})^{2}(w_j^{h})^{2}+\alpha\sum_{\substack{h'=1\\h'\neq h}}^{s}(v_h)^{\beta}\sum_{i=1}^{n}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^{2}(w_j^{h})^{2}}$$

Proof By taking the derivative of the Co-FW-MVFCM objective function w.r.t a_{kj}^{h} and setting them to be zero, we obtain the equation

Co-FW-MVFCM: Shortened Proof

$$
\begin{aligned}
\frac{\partial J}{\partial a_{kj}^{h}} &= (v_h)^{\beta}\Bigg(\Big\{-2\sum_{i=1}^{n}(\mu_{ik}^{h})^2(w_j^h)^2(x_{ij}^h-a_{kj}^h)\Big\} \\
&\quad +\Big\{-2\alpha\sum_{h'=1,h'\neq h}^{s}\sum_{i=1}^{n}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^2(w_j^h)^2(x_{ij}^h-a_{kj}^h)\Big\}\Bigg)=0 \\
\Rightarrow &\quad a_{kj}^{h}\Big((v_h)^{\beta}\sum_{i=1}^{n}(\mu_{ik}^{h})^2(w_j^h)^2+\alpha\sum_{h'=1,h'\neq h}^{s}(v_h)^{\beta}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^2(w_j^h)^2\Big)= \\
&\quad (v_h)^{\beta}\sum_{i=1}^{n}(\mu_{ik}^{h})^2(w_j^h)^2x_{ij}^h+\alpha\sum_{h'=1,h'\neq h}^{s}(v_h)^{\beta}\{\mu_{ik}^{h}-\mu_{ik}^{h'}\}^2(w_j^h)^2x_{ij}^h
\end{aligned}
$$

Thus, the updating equation for a_{kj}^{h} can be obtained as Equation (38)

Co-FW-MVFCM: Parameter Estimations

3 Parameters

1. Threshold
2. Exponent parameter
3. Collaborating parameter

Why are conditionally parameters estimation necessary?

中原大學 Chung Yuan Christian University

Co-FW-MVFCM: Parameter Estimations

We have the threshold equation!

$$W^{h^t} < 1 \Big/ \sqrt{nd_h} \tag{43}$$

What about β?

$$\beta = \frac{t}{s} \tag{44}$$

and α?

$$\alpha = \frac{t}{n} \tag{45}$$

which is the effectiveness of these 3 parameters, confirmed by experiment.

中原大學 Chung Yuan Christian University

How Co-FW-MVFCM's without Feature-reduction Algorithm Works?

Co-FW-MVFCM without Feature-reduction Algorithm

So how does a Co-FW-MVFCM without Feature-reduction algorithm work, I will explain this using seven steps:

Input Data set $\boldsymbol{X} = \{x_1, ..., x_n\}$ with $x_i = \{x_i^h\}_{h=1}^s$ and $x_i^h = \{x_{ij}^h\}_{j=1}^{d_h}$, number c of cluster, $\alpha = 0$, $\beta = 0$, and $\varepsilon > 0$

Output $w_j^h, v_h, a_{kj}^h, \mu_{ik}^h$, and $\hat{\mu_{ik}}$

Initialization Randomly generate initial U^h, Cluster centers A^h, initialize feature weight $W^{h^t} = [w_j^h]_{1\times d_h}$ (user may define $w_j^h = 1/d_h \forall j$), initialize view weight $V^t = [v_h]_{1\times s}$ (user may define $v_h = 1/s \forall h$), and set $t = 1$.

Co-FW-MVFCM without Feature-reduction Algorithm

Step 1 Calculate α and β by $\beta = t/n$.

Step 2 Udate the feature weight W^{h^t} using $A^{h^{t-1}}, U^{h^{t-1}}$, and V^{t-1} by Eq. (37) .

Step 3 Update the view weight V^{h^t} using $A^{h^{t-1}}, U^{h^{t-1}}$, and W^{h^t} by Eq. (36).

Step 4 Update the cluster centers A^{h^t} using W^{h^t}, V^{h^t}, and $U^{h^{t-1}}$ by Eq. (38).

Step 5 Update membership U^{h^t} using W^{h^t}, A^{h^t} and α by Eq. (35).

中原大學 Chung Yuan Christian University

Co-FW-MVFCM without Feature-reduction Algorithm

Step 6 If $\left(\left|\|U^t\| - \|U^{t-1}\|\right|\right) \big/ nc < \varepsilon$, then stop;
Else set $t = t + 1$ and go back to Step 1.

Step 7 Compute the global fuzzy partition matrix by Eq. (34).

中原大學 Chung Yuan Christian University

How Co-FW-MVFCM's with Feature-reduction Algorithm Works?

中原大學 Chung Yuan Christian University

Co-FW-MVFCM with Feature-reduction Algorithm

So how does a Co-FW-MVFCM with Feature-reduction algorithm work, I will explain this using nine steps:

Input Data set $\boldsymbol{X} = \{x_1, ..., x_n\}$ with $x_i = \{x_i^h\}_{h=1}^s$ and $x_i^h = \{x_{ij}^h\}_{j=1}^{d_h}$, number c of cluster, $\alpha = 0$, $\beta = 0$, and $\varepsilon > 0$

Output $w_j^h, v_h, a_{kj}^h, \mu_{ik}^h$, and $\hat{\mu_{ik}}$

Initialization Randomly generate initial U^h, Cluster centers A^h, initialize feature weight $W^{h^t} = [w_j^h]_{1\times d_h}$ (user may define $w_j^h = 1/d_h \forall j$), initialize view weight $V^t = [v_h]_{1\times s}$ (user may define $v_h = 1/s\forall h$), and set $t = 1$.

Co-FW-MVFCM with Feature-reduction Algorithm

Step 1 Calculate α and β by Eqs. (45) and (44)

Step 2 Update the feature weight W^{h^t} using $A^{h^{t-1}}, U^{h^{t-1}}$, and V^{t-1} by Eq. (37) .

Step 3 Discard total d_r number of these j feature components for $W^{h^t} < 1/\sqrt{nd_h}$, and set $d = d^{new} = D - d_r$.

Step 4 Adjust $W^{h^t}$by $(w_j^h)' = w_j^h \Big/ \sum_{p=1}^{d_h(new)} w_p^h$.

中原大學 Chung Yuan Christian University

Co-FW-MVFCM with Feature-reduction Algorithm

Step 5 Update the view weight V^{h^t} using $A^{h^{t-1}}, U^{h^{t-1}}$, and W^{h^t} by Eq. (36).

Step 6 Update the cluster centers A^{h^t} using W^{h^t}, V^{h^t}, and $U^{h^{t-1}}$ by Eq. (38).

Step 7 Update membership U^{h^t} using W^{h^t}, A^{h^t} and α by Eq. (35).

Step 8 If $\left(\left|\|U^t\| - \|U^{t-1}\|\right|\right) / nc < \varepsilon$, then stop;
Else set $t = t + 1$ and go back to Step 1.

Step 9 Compute the global fuzzy partition matrix by Eq. (34).

中原大學 Chung Yuan Christian University

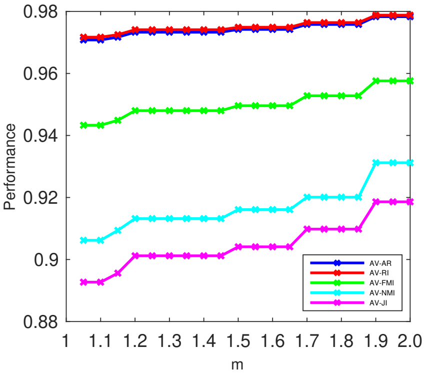
Experimental Comparisons and Results

Parameter selections

Algorithms	Data set	Parameter setup: m	β	η	λ_h
Co-FKM	Prokaryotic	$m=\frac{\min(n,d-1)}{\max(n,0.5(d-s)-s^2C(s+c,c-1))}$	-	$\eta=\frac{s-1}{s}$	-
	MSRC-V1				
	Caltech 101-7	m = [1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50]		$\eta=0.3$	
	IS			$\eta=\frac{s-1}{s}$	
	IA				
	Wikipedia articles	m = [1.05 1.10 1.15 1.25 1.35 1.5 1.75 2]			
MultiNMF	Prokaryotic	-	-	-	0.01
	MSRC-V1				
	Caltech 101-7				
	IS				
	IA				
	Wikipedia articles				
WV-Co-FCM	Prokaryotic	$m=\frac{\min(n,d-1)}{\max(n,0.5(d-s)-s^2C(s+c,c-1))}$	$\beta>0$	$\eta=\frac{s-1}{s}$	-
	MSRC-V1				
	Caltech 101-7	m = [1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50]			
	IS				
	IA				
	Wikipedia articles	m = [1.05 1.10 1.15 1.25 1.35 1.5 1.75 2]			
Minimax-FCM	Prokaryotic	$m=\frac{\min(n,d-1)}{\max(n,0.5(d-s)-s^2C(s+c,c-1))}$	$\beta=\frac{s-1}{s}$	-	-
	MSRC-V1				
	Caltech 101-7	m = [1.05 1.10 1.15 1.20 1.25 1.30 1.35 1.40 1.45 1.50]			
	IS				
	IA				
	Wikipedia articles	m = [1.05 1.10 1.15 1.25 1.35 1.5 1.75 2]			

Result: Analysis of the incorporated idea

- Performance comparison of m ranges from 1.05 to 2 with step 0.05 for a two-view numerical data set with 4 clusters and 3 feature components on FW-MVFCM



Result: Analysis of the incorporated idea

- Equal initial feature-view weights, AR, RI, FMI, NMI, and JI for different iterations of a two-view numerical data set with 4 clusters and 3 feature components

Iteration No.	beta	Feature weights						FW-MVFCM				
		$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	AR	RI	FMI	NMI	JI
0	0	0.1667	0.3333	0.5000	0.1667	0.3333	0.5000	-	-	-	-	-
1	8.33E-04	0.3723	0.6276	0.0001	0.5151	0.4847	0.0002	0.4863	0.6697	0.5409	0.3475	0.3514
2	0.0017	0.3721	0.6278	0.0001	0.5158	0.4840	0.0002	0.5267	0.7234	0.4983	0.3736	0.3300
3	0.0025	0.3716	0.6283	0.0001	0.5165	0.4833	0.0002	0.5300	0.7632	0.5332	0.4395	0.3635
4	0.0033	0.3715	0.6284	0.0001	0.5169	0.4829	0.0003	0.6817	0.7750	0.5510	0.4633	0.3803
5	0.0042	0.3701	0.6298	0.0001	0.5158	0.4839	0.0003	0.7300	0.7918	0.6058	0.5231	0.4336
6	0.005	0.3637	0.6362	0.0001	0.5112	0.4885	0.0003	0.7342	0.7941	0.6191	0.5468	0.4466
7	0.0058	0.3499	0.6500	0.0001	0.5012	0.4985	0.0003	0.7625	0.8120	0.6473	0.5718	0.4772
8	0.0067	0.3346	0.6652	0.0001	0.4902	0.5094	0.0004	0.8308	0.8600	0.7308	0.6543	0.5751
9	0.0075	0.3469	0.6529	0.0002	0.4934	0.5061	0.0005	0.8900	0.9034	0.8098	0.7419	0.6802
10	0.0083	0.3700	0.6295	0.0005	0.5069	0.4921	0.0010	0.9508	0.9538	0.9076	0.8623	0.8309
11	0.0092	0.3628	0.6364	0.0008	0.4891	0.5096	0.0014	0.9725	0.9733	0.9465	0.9134	0.8984
12	0.01	0.3701	0.6290	0.0009	0.4871	0.5115	0.0014	0.9775	0.9781	0.9560	0.9295	0.9157
13	0.0108	0.3704	0.6287	0.0009	0.4864	0.5122	0.0014	0.9783	0.9788	0.9576	0.9312	0.9186
14	0.0117	0.3707	0.6284	0.0009	0.4864	0.5122	0.0014	0.9783	0.9788	0.9576	0.9312	0.9186

Result: Analysis of the incorporated idea

- Different initial feature-view weights, AR, RI, FMI, NMI, and JI for different iterations of a two-view numerical data set with 4 clusters and 3 feature components

Iteration No.	Feature weights						FW-MVFCM				
	$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	AR	RI	FMI	NMI	JI
0	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	-	-	-	-	-
1	0.3719	0.6280	0.0001	0.5145	0.4853	0.0002	0.5000	0.7223	0.4763	0.3556	0.3119
2	0.3720	0.6279	0.0001	0.5145	0.4853	0.0002	0.4858	0.7259	0.4644	0.3529	0.4644
3	0.3715	0.6284	0.0001	0.5138	0.4860	0.0002	0.5383	0.7462	0.5314	0.4210	0.5314
4	0.3694	0.6305	0.0001	0.5110	0.4887	0.0003	0.5592	0.7487	0.5630	0.4566	0.5630
5	0.3613	0.6386	0.0001	0.5033	0.4964	0.0003	0.5742	0.7517	0.5878	0.4876	0.5878
6	0.3469	0.6530	0.0001	0.4833	0.5114	0.0003	0.5858	0.7586	0.5927	0.4927	0.5927
7	0.3283	0.6715	0.0001	0.4682	0.5315	0.0003	0.6050	0.7792	0.5990	0.5131	0.5990
8	0.3206	0.6793	0.0002	0.4509	0.5487	0.0004	0.6800	0.8216	0.6720	0.5847	0.6720
9	0.3415	0.6583	0.0002	0.4531	0.5464	0.0005	0.6892	0.8432	0.7933	0.6421	0.7033
10	0.3633	0.6364	0.0003	0.4789	0.5204	0.0007	0.8983	0.9092	0.8196	0.7577	0.8196
11	0.3595	0.6401	0.0005	0.4969	0.5020	0.0011	0.9583	0.9741	0.9201	0.8766	0.9201
12	0.3642	0.6350	0.0008	0.4873	0.5113	0.0014	0.9733	0.9601	0.9481	0.9155	0.9481
13	0.3698	0.6293	0.0009	0.4871	0.5114	0.0014	0.9775	0.9781	0.9560	0.9295	0.9560
14	0.3704	0.6287	0.0009	0.4864	0.5122	0.0014	0.9783	0.9788	0.9576	0.9312	0.9576
15	0.3707	0.6284	0.0009	0.4864	0.5122	0.0014	0.9783	0.9788	0.9576	0.9312	0.9576
16	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.9783	0.9788	0.9576	0.9312	0.9576
17	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.9783	0.9788	0.9576	0.9312	0.9576

Result:

- Different initial feature-view weights, AR, RI, FMI, NMI, and JI of Co-FW-MVFCM without feature reduction (FR) for different iterations of a two-view numerical data set with 4 clusters and 3 feature components

Iteration No.	Feature weights						Co-FW-MVFCM without FR				
	$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	AR	RI	FMI	NMI	JI
0	0.4942	0.1674	0.3383	0.1440	0.2598	0.5962	-	-	-	-	-
1	0.3538	0.6394	0.0068	0.4426	0.5490	0.0084	0.4842	0.6564	0.5233	0.3618	0.3358
2	0.3717	0.6281	0.0001	0.5138	0.4860	0.0002	0.4833	0.6852	0.5145	0.3722	0.3359
3	0.3688	0.6310	0.0001	0.5115	0.4883	0.0002	0.6242	0.7854	0.6552	0.6573	0.4762
4	0.3287	0.6712	0.0001	0.5059	0.4939	0.0002	0.9975	0.9975	0.9950	0.9879	0.9901
5	0.3137	0.6852	0.0001	0.4685	0.5313	0.0002	1	1	1	1	1

Result:

- Different initial feature-view weights, AR, RI, FMI, NMI, and JI of Co-FW-MVFCM with feature reduction (FR) for different iterations of a two-view numerical data set with 4 clusters and 3 feature components

Iteration No.	Feature weights						Co-FW-MVFCM with FR				
	$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	AR	RI	FMI	NMI	JI
Init	0.4942	0.1674	0.3383	0.1440	0.2598	0.5962	-	-	-	-	-
0	0.3538	0.6394	0.0068	0.4426	0.5490	0.0084	-	-	-	-	-
1	0.3562	0.6438	-	0.4463	0.5537	-	0.4842	0.6564	0.5233	0.3618	0.3358
2	0.3718	0.6282	-	0.5139	0.4861	-	0.4833	0.6852	0.5145	0.3722	0.3359
3	0.3689	0.6311	-	0.5116	0.4884	-	0.6250	0.7858	0.6556	0.6577	0.4766
4	0.3287	0.6713	-	0.5060	0.4940	-	0.9975	0.9975	0.9950	0.9879	0.9901
5	0.3148	0.6852	-	0.4687	0.5313	-	1	1	1	1	1

中原大學 Chung Yuan Christian University

Result:

- Final feature-view weights of FW-MVFCM and Co-FW-MVFCM with feature reduction (FR) of a two-view numerical data set with 4 clusters and 3 feature Components over 5 simulations

Simulation	Final feature-view weights from FW -MVFCM with *m=2*						Final feature-view weights from Co-FW-MVFCM with FR					
	$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$
1	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.3684	0.6316	-	0.5863	0.4137	-
2	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.3694	0.6306	-	0.4548	0.5452	-
3	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.3669	0.6331	-	0.5925	0.4075	-
4	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.3705	0.6295	-	0.5142	0.4858	-
5	0.3707	0.6284	0.0009	0.4863	0.5122	0.0014	0.3148	0.6852	-	0.4687	0.5313	-

Result:

- Comparison of AR, RI, FMI, NMI, and JI between FW-MVFCM and Co-FW-MVFCM based on the simulations of Table 5.7

	FW-MVFCM with *m=2*					Co-FW-MVFCM with FR				
	Sim. 1	Sim. 2	Sim. 3	Sim. 4	Sim. 5	Sim. 1	Sim. 2	Sim. 3	Sim. 4	Sim. 5
AR	0.9783	0.9783	0.9783	0.9783	0.9783	0.9950	0.9900	0.9983	0.9933	1
RI	0.9788	0.9788	0.9788	0.9788	0.9788	0.9951	0.9900	0.9983	0.9934	1
FMI	0.9576	0.9576	0.9576	0.9576	0.9576	0.9901	0.9800	0.9967	0.9867	1
NMI	0.9312	0.9312	0.9312	0.9312	0.9312	0.9787	0.9596	0.9919	0.9717	1
JI	0.9186	0.9186	0.9186	0.9186	0.9186	0.9804	0.9608	0.9956	0.9737	1

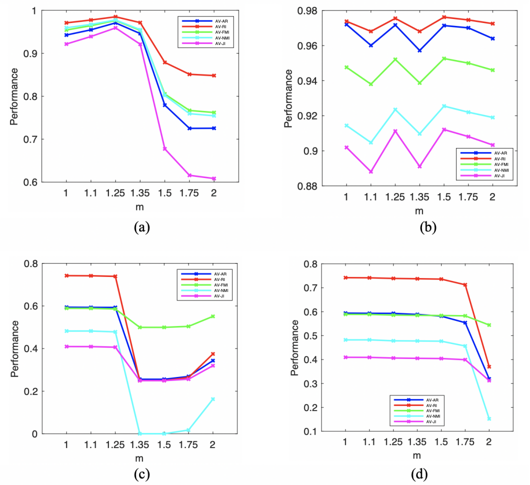
中原大學 Chung Yuan Christian University

Result:

- Performance comparison of m are [1.05 1.10 1.25 1.35 1.50 1.75 2] for a two-view numerical data set with 4 clusters and 3 feature components on (a) Co-FKM, (b) MinMax FCM, (c) WV-Co-FCM for the first case of δ_{ik}^{h} with $\delta_{ik}^{h} = \eta \frac{1}{s-1} \sum_{h'=1, h' \neq h}^{s} \sum_{j=1}^{d_h} \left(x_{ij}^{h'} - a_{kj}^{h'} \right)^2$, and (d) WV-Co-FCM for the second case of δ_{ik}^{h} with $\delta_{ik}^{h} = \frac{\eta}{s} \sum_{h=1}^{s} \sum_{j=1}^{d_h} \left(x_{ij}^{h} - a_{kj}^{h} \right)^2$

中原大學 Chung Yuan Christian University

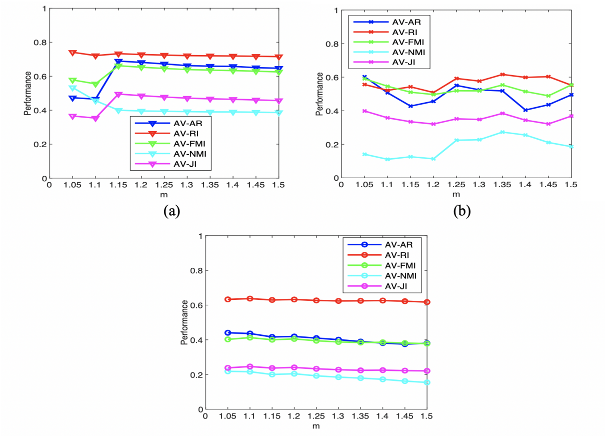
Result:



(a) (b) (c) (d)

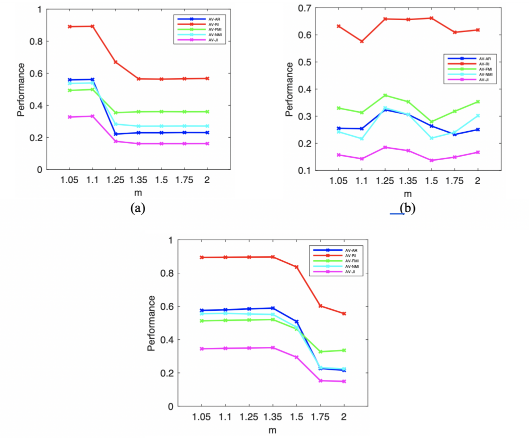
Result:

- Performance comparison of m ranges from 1.05 to 1.5 and $\eta = s - 1/s$ for (a) Co-FKM; (b) WV-Co-FCM; and (c) Minimax-FCM on Caltech 101 data set.



Result:

- Performance comparison of m ranges from 1.05 to 2 and $\eta = s - 1/s$ for (a) Co-FKM; (b) WV-Co-FCM; and (c) Minimax-FCM on Wikipedia Articles data set.



中原大學 Chung Yuan Christian University

Result:

- Classification results of CoFKM, MultiNMF, WV-Co-FCM, Minimax FCM, and the proposed Co-FW-MVFCM algorithms in Prokaryotic Phyla and MSRC-V124 multi-view data sets

		Co-FKM	MultiNMF	WV-Co-FCM		Minimax-FCM	Co-FW-MVFCM
				The 1st case of δ_{ik}^{h}	The 2nd case of δ_{ik}^{h}		
Prokaryotic Phyla	AR	0.4410/0.5381/0.5408	0.2759/0.4328/0.4773	0.5009/0.5178/0.5227	0.5644/0.5674/0.5681	0..4483/0.4483/0.4483	0.4482/**0.6306**/0.6842
	RI	0.6090/0.6117/0.6360	0.5656/0.5914/0.6046	0.5602/0.5644/0.5667	0.3933/0.3970/0.3996	0.5970/0.5970/0.5970	0.5967/**0.6553**/0.7058
	FMI	0.4518/0.4539/0.4586	0.3703/0.3953/0.4065	0.4961/0.5145/0.5182	0.6236/**0.6277**/0.6287	0.3856/0.3856/0.3856	0.4464/0.5349/0.5897
	NMI	0.3640/**0.3883**/0.3910	0.1940/0.2161/0.2246	0.2370/0.2427/0.2475	0.0000/0.0290/0.0413	0.2061/0.2061/0.2061	0.2381/0.3056/0.3461
	JI	0.2985/0.2910/0.2911	0.2201/0.2412/0.2512	0.3281/0.3433/0.3464	0.3913/0.3946/0.3958	0.2316/0.2316/0.2316	0.2860/**0.3640**/0.4147
MSRC-V1	AR	0.5952/0.6523/0.7429	0.4190/0.5625/0.6810	0.1714/0.2494/0.3143	0.1381/0.1915/0.2667	0.2381/0.2743/0.3095	0.6143/**0.6576**/0.7571
	RI	0.8458/0.8664/0.8854	0.8099/0.8411/0.8703	0.5624/0.6694/0.7465	0.5548/0.5779/0.7674	0.7815/0.7846/0.7855	0.8491/**0.8687**/0.8938
	FMI	0.5440/**0.5791**/0.6121	0.3908/0.4662/0.5545	0.2414/0.2727/0.3012	0.2035/0.2801/0.2926	0.2325/0.2420/0.2440	0.4597/0.5351/0.6222
	NMI	0.5550/**0.5934**/0.6346	0.4170/0.4978/0.5856	0.1079/0.2494/0.3143	0.1198/0.1416/0.1855	0.2284/0.2363/0.2388	0.4785/0.5527/0.6297
	JI	0.3666/**0.4036**/0.4400	0.2426/0.3039/0.3830	0.1335/0.1446/0.1563	0.1131/0.1377/0.1484	0.1315/0.1376/0.1389	0.2985/0.3661/0.4515

中原大學 Chung Yuan Christian University

Result:

- Classification results of CoFKM, MultiNMF, WV-Co-FCM, Minimax FCM, and the proposed Co-FW-MVFCM algorithms in Caltech 101-7, Image segmentation (IS), and Internet advertisement (IA) multi-view data sets

		Co-FKM	MultiNMF	WV-Co-FCM	Minimax-FCM	Co-FW-MVFCM
Caltech 101-7	AR	0.3704/0.6258/0.7564	0.3403/0.3860/0.4817	0.2714/0.4918/0.7463	0.3304/0.4050/0.4586	0.6425/**0.6795**/0.7225
	RI	0.6208/0.7239/0.7873	0.6574/0.6910/0.7320	0.3911/0.5664/0.7502	0.5978/0.6273/0.6554	0.6942/**0.7450**/0.7774
	FMI	0.4429/0.6256/0.7318	0.4423/0.4755/0.5615	0.4024/0.5288/0.6909	0.3557/0.3932/0.4325	0.6122/**0.6576**/0.6979
	NMI	0.3245/0.4123/0.5849	0.3833/**0.4330**/0.4959	0.0085/0.1864/0.4469	0.1375/0.1885/0.2645	0.3054/0.3521/0.4000
	JI	0.2807/0.4496/0.5736	0.2460/0.2772/0.3492	0.2476/0.3525/0.5270	0.1966/0.2317/0.2646	0.4409/**0.4888**/0.5350
IS	AR	0.3623/0.5139/0.6398	0.4745/0.5596/0.6190	0.1203/0.2453/0.4563	0.2835/0.5044/0.5169	0.5502/**0.5794**/0.6463
	RI	0.7887/0.8360/0.8562	0.8312/0.8568/0.8794	0.1425/0.6105/0.7857	0.7754/0.8423/0.8498	0.8382/**0.8575**/0.8711
	FMI	0.3631/0.4700/0.5319	0.4634/**0.5260**/0.5965	0.1987/0.3095/0.5287	0.3889/0.5090/0.5250	0.4790/0.5157/0.5513
	NMI	0.4119/0.5302/0.5976	0.5262/**0.5743**/0.6293	0.0000/0.1601/0.5682	0.4385/0.5609/0.5727	0.4800/0.5256/0.5714
	JI	0.2215/0.3066/0.3611	0.3000/**0.3564**/0.4224	0.1072/0.1603/0.3053	0.2373/0.3391/0.3538	0.3138/0.3473/0.3805
IA	AR	0.7067/0.8537/0.8957	0.8381/0.8382/0.8385	0.8025/0.8329/0.8368	0.8283/0.8381/0.8385	0.8631/**0.8930**/0.8949
	RI	0.5852/0.7543/0.8131	0.7331/0.7335/0.7354	0.7127/0.7445/0.7490	0.7155/0.7456/0.7460	0.7636/**0.8088**/0.8118
	FMI	0.6786/0.8264/0.8762	0.8550/0.8552/0.8561	0.8000/0.8560/0.8620	0.8426/0.8596/0.8598	0.8335/**0.8728**/0.8752
	NMI	0.1117/0.2406/0.3059	0.0471/0.0513/0.747	0.0289/0.1052/0.1245	0.0000/0.1219/0.1240	0.2476/**0.2978**/0.3028
	JI	0.5035/0.7055/0.7791	0.7315/0.7318/0.7333	0.6892/0.7366/0.7419	0.7145/0.7401/0.7404	0.7140/**0.7738**/0.7775

中原大學 Chung Yuan Christian University

Result:

- Classification results of MultiNMF and the proposed Co-FW-MVFCM algorithms in Wikipedia Articles multi-view data set

	AR	RI	FMI	NMI	JI
MultiNMF	0.4791/0.5254/0.5642	0.8728/0.8829/0.8898	0.4334/0.4599/0.4860	0.4891/0.5167/0.5366	0.2766/0.2987/0.3210
Co-FW-MVFCM	0.5902/**0.5932**/0.5974	0.8935/**0.8960**/0.8978	0.5071/**0.5169**/0.5259	0.5459/**0.5517**/0.5609	0.3397/**0.3485**/0.3567

- Results of final features of the proposed Co-FW-MVFCM

Data sets	Original *d*						Final *d* by Co-FW-MVFCM					
	v_1	v_2	v_3	v_4	v_5	v_6	v_1	v_2	v_3	v_4	v_5	v_6
Prokaryoric Phyla	393	3	438	-	-	-	48	2	139	-	-	-
MSRC V-1	24	576	512	256	254	-	24	27	32	64	68	-
Caltech 101-7	48	40	254	1984	512	928	48	40	244	461	377	466
Image segmentation (IS)	9	10	-	-	-	-	2	8	-	-	-	-
Internet advertisement (IA)	3	457	495	472	111	19	2	270	288	270	100	12
Wikipedia Articles	128	10	-	-	-	-	85	10	-	-	-	-

Conclusions

Conclusions

- According to the experimental results, the proposed FW-MVFCM is capable of selecting the un-informative features in one view by consistently weighting the un-informative feature with smaller values.
- FW-MVFCM does not enforce any assumptions on the form of correlation between views.
- The collaborative framework in the proposed Co-FW-MVFCM contains a two-step procedure that is a local step and a collaborative step.
- The proposed Co-FW-MVFCM algorithm with a feature reduction step can exclude redundant features in one view if those features are smaller than the threshold.

C.J. Merz [27] C.L. Blake. "UCI repository of machine learning databases, a huge collection of artificial and real-world data sets". In: (1998).

Witold Pedrycz. "Collaborative fuzzy clustering". In: *Pattern Recognition Letters* (2002). ISSN: 01678655. DOI: 10.1016/S0167-8655(02)00130-7.

J. Winn and N. Jojic. "LOCUS: Learning Object Classes with Unsupervised Segmentation". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2005. ISBN: 076952334X. DOI: 10.1109/ICCV.2005.148.

中原大學 Chung Yuan Christian University

Li Fei-Fei, Rob Fergus, and Pietro Perona. "Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories". In: *Computer Vision and Image Understanding* (2007). ISSN: 10773142. DOI: 10.1016/j.cviu.2005.09.012.

Guillaume Cleuziou et al. "CoFKM: A centralized method for multiple-view clustering". In: *Proceedings - IEEE International Conference on Data Mining, ICDM*. 2009. ISBN: 9780769538952. DOI: 10.1109/ICDM.2009.138.

Jialu Liu et al. "Multi-view clustering via joint nonnegative matrix factorization". In: *Proceedings of the 2013 SIAM International Conference on Data Mining, SDM 2013*. 2013. ISBN: 9781611972627. DOI: 10.1137/1.9781611972832.28.

Jose Costa Pereira et al. "On the role of correlation and abstraction in cross-modal multimedia retrieval". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2014). ISSN: 01628828. DOI: 10.1109/TPAMI.2013.142.

Yizhang Jiang et al. "Collaborative Fuzzy Clustering From Multiple Weighted Views". In: *IEEE Transactions on Cybernetics* (2015). ISSN: 21682267. DOI: 10.1109/TCYB.2014.2334595.

Maria Brbić et al. "The landscape of microbial phenotypic traits and associated genes". In: *Nucleic Acids Research* (2016). ISSN: 13624962. DOI: 10.1093/nar/gkw964.

Yangtao Wang and Lihui Chen. "Multi-view fuzzy clustering with minimax optimization for effective clustering of data from multiple sources". In: *Expert Systems with Applications* (2017). ISSN: 09574174. DOI: 10.1016/j.eswa.2016.10.006.


THANK
YOU